Amendments to the Specification

On page 15, the 3rd paragraph, beginning at line 26, is amended to read as follows:

Alignment can be used to measure the level of fitness between a kernel and a fixed labeling of the data, with the goal of selecting better kernel parameters, i.e., the best aligned set of labels. The absolute measure of a kernel is its second eigenvalue, which can be used to optimize kernel parameters. Using eigentechniques, the measure of the kernel can be lower-bounded with the second eigenvalue.

On page 15, before the 4th paragraph, insert the following new paragraphs:

Consider the constraint C1: $\sum_{+} p_i = \sum_{-} p_i$. The requirement that y = -1 or +1 is relaxed and the following conditions are applied: $\sum y_i^2 = m$ and $\sum_{+} p_i = \sum_{-} p_i$, so that the two classes have equal probability under the distribution of the first eigenvector.

Under these constraints, the alignment can be maximized by spectral techniques. Defining $v = \frac{y}{\sqrt{(m)}}$, the problem becomes:

$\min_{C1, \sum y_i^2 - m} y'Ky = \min_{\sum v_2^2 = 1; + otherconstra\text{int}} \frac{v'Kv}{v'v} = \lambda_2$, where v_2 is the minimizer, the second eigenvector. Constraint C1 requires the probability of the two classes to be the same under the distribution given by the first eigenvector.

In other words, the second eigenvector maximizes the alignment under the constraint, and thresholding it provides a labeling that approximately maximizes alignment. The second eigenvalue gives the value of the alignment in this case, and is a lower bound under the true value of the optimal alignment. This provides an absolute measure of kernel alignment so that the maximal alignment can be achieved on all possible labeling. One can also tune the kernel parameters in a principled way, possibly by gradient descent, to achieve maximal alignment.

On page 22, the 4th paragraph, beginning at line 22, is amended to read as follows:

Using the breast cancer dataset with both linear and Gaussian kernel, cut cost (see Equation 5) was used to select the best threshold. The results are plotted in FIGS. 4(a) and (b), where cut cost

$C(S, k, y)$-$\lambda/(2\|K\|_F$ for y=sign(v^{max}- θ_i) is indicated by the dashed ~~curved~~ curve and error rate of y against threshold i is indicated by the solid curve. For the linear kernel, FIG. 4(a), the threshold was set at 378 with an accuracy of 67.86%, which is significantly worse ~~that~~ than the results obtained by optimizing the alignment. With the Gaussian kernel, FIG. 4(b), the method selects threshold 312 with an accuracy of 80.31%, a slight improvement over the results obtained with the Gaussian kernel by optimizing the alignment.